METHOD OF INDEXED STORAGE AND RETRIEVAL OF MULTIDIMENSIONAL INFORMATION

ABSTRACT OF THE DISCLOSURE

A tree-structured index to multidimensional data is created using naturally occurring patterns and clusters within the data which permit efficient search and retrieval strategies in a database of DNA profiles. A search engine utilizes hierarchical decomposition of the database by identifying clusters of similar DNA profiles and maps to parallel computer architecture, allowing scale up past previously feasible limits. Key benefits of the new method are logarithmic scale up and parallelization. These benefits are achieved by identification and utilization of naturally occurring patterns and clusters within stored data. The patterns and clusters enable the stored data to be partitioned into subsets of roughly equal size. The method can be applied recursively, resulting in a database tree that is balanced, meaning that all paths or branches through the tree have roughly the same length. The method achieves high performance by exploiting the natural structure of the data in a manner that maintains balanced trees. Implementation of the method maps naturally to parallel computer architectures, allowing scale up to very large databases.